



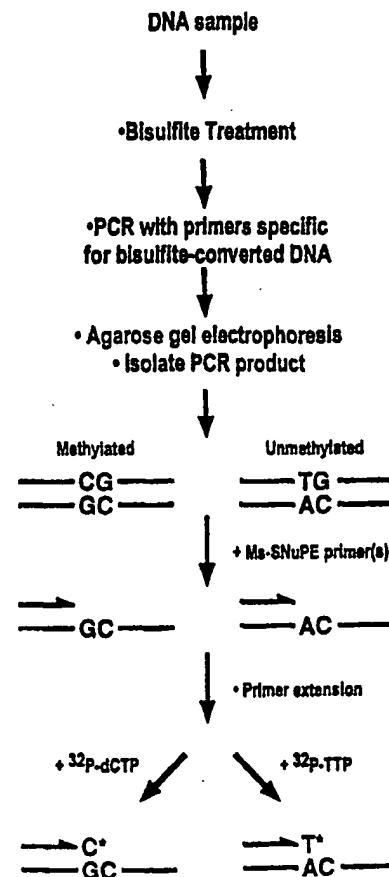
## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>C12Q 1/68, C07H 21/04</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 98/56952</b> <b>(43) International Publication Date:</b> 17 December 1998 (17.12.98)
<b>(21) International Application Number:</b> PCT/US98/11896 <b>(22) International Filing Date:</b> 9 June 1998 (09.06.98) <b>(30) Priority Data:</b> 60/049,231      9 June 1997 (09.06.97)      US <b>(71) Applicant:</b> UNIVERSITY OF SOUTHERN CALIFORNIA [US/US]; University Park, Los Angeles, CA 90089 (US). <b>(72) Inventors:</b> GONZALGO, Mark, L.; 1218 Pine Street #304, South Pasadena, CA 91030 (US). JONES, Peter, A.; 4645 Lasheart Drive, La Canada, CA 91011 (US). LIANG, Ganging; 3436 Ashbourne Place, Rowland Heights, CA 91748 (US). <b>(74) Agent:</b> OSTER, Jeffrey, B.; Davis Wright Tremaine, 2600 Cen- tury Square, 1501 Fourth Avenue, Seattle, WA 98101-1688 (US).		<b>(81) Designated States:</b> AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KR, KZ, LT, LU, LV, MD, MK, MN, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, UZ, VN, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i>

**(54) Title:** A CANCER DIAGNOSTIC METHOD BASED UPON DNA METHYLATION DIFFERENCES

**(57) Abstract**

There is disclosed a cancer diagnostic method based upon DNA methylation differences at specific CpG sites. As set forth in the Figure, the method comprises bisulfite treatment of DNA, followed by methylation-sensitive single nucleotide primer extension (Ms-SNuPE) for determination of strand-specific methylation status at cytosine residues.



**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## A CANCER DIAGNOSTIC METHOD BASED UPON DNA METHYLATION DIFFERENCES

### 5    **Technical Field of the Invention**

          The present invention provides a cancer diagnostic method based upon DNA methylation differences at specific CpG sites. Specifically, the inventive method provides for a bisulfite treatment of DNA, followed by methylation-sensitive single nucleotide primer extension (Ms-SNuPE), for determination of strand-specific methylation status at cytosine  
10    residues.

### **Background of the Invention**

          Cancer treatments, in general, have a higher rate of success if the cancer is diagnosed early and treatment is started earlier in the disease process. The relationship between improved  
15    prognosis and stage of disease at diagnosis hold across all forms of cancer for the most part. Therefore, there is an important need to develop early assays of general tumorigenesis through marker assays that measure general tumorigenesis without regard to the tissue source or cell type that is the source of a primary tumor. Moreover, there is a need to address distinct genetic alteration patterns that can serve as a platform associated with general tumorigenesis for early  
20    detection and prognostic monitoring of many forms of cancer.

#### Importance of DNA Methylation

          DNA methylation is a mechanism for changing the base sequence of DNA without altering its coding function. DNA methylation is a heritable, reversible and epigenetic change. Yet, DNA methylation has the potential to alter gene expression, which has profound  
25    developmental and genetic consequences. The methylation reaction involves flipping a target cytosine out of an intact double helix to allow the transfer of a methyl group from S-adenosylmethionine in a cleft of the enzyme DNA (cystosine-5)-methyltransferase (Klimasauskas et al., *Cell* 76:357-369, 1994) to form 5-methylcytosine (5-mCyt). This enzymatic conversion is the only epigenetic modification of DNA known to exist in vertebrates  
30    and is essential for normal embryonic development (Bird, *Cell* 70:5-8, 1992; Laird and Jaenisch, *Human Mol. Genet.* 3:1487-1495, 1994; and Bestor and Jaenisch, *Cell* 69:915-926, 1992). The presence of 5-mCyt at CpG dinucleotides has resulted in a 5-fold depletion of this sequence in the genome during vertebrate evolution, presumably due to spontaneous deamination of 5-mCyt to T (Schoreret et al., *Proc. Natl. Acad. Sci. USA* 89:957-961, 1992).  
35    Those areas of the genome that do not show such suppression are referred to as "CpG islands" (Bird, *Nature* 321:209-213, 1986; and Gardiner-Garden et al., *J. Mol. Biol.* 196:261-282, 1987). These CpG island regions comprise about 1% of vertebrate genomes and also account for about 15% of the total number of CpG dinucleotides (Bird, *Infra.*). CpG islands are typically between 0.2 to about 1 kb in length and are located upstream of many housekeeping

and tissue-specific genes, but may also extend into gene coding regions. Therefore, it is the methylation of cytosine residues within CpG islands in somatic tissues, which is believed to affect gene function by altering transcription (Cedar, *Cell* 53:3-4, 1988).

5 Methylation of cytosine residues contained within CpG islands of certain genes has been inversely correlated with gene activity. This could lead to decreased gene expression by a variety of mechanisms including, for example, disruption of local chromatin structure, inhibition of transcription factor-DNA binding, or by recruitment of proteins which interact specifically with methylated sequences indirectly preventing transcription factor binding. In other words, there are several theories as to how methylation affects mRNA transcription and  
10 gene expression, but the exact mechanism of action is not well understood. Some studies have demonstrated an inverse correlation between methylation of CpG islands and gene expression, however, most CpG islands on autosomal genes remain unmethylated in the germline and methylation of these islands is usually independent of gene expression. Tissue-specific genes are usually unmethylated and the receptive target organs but are methylated in the germline and  
15 in non-expressing adult tissues. CpG islands of constitutively-expressed housekeeping genes are normally unmethylated in the germline and in somatic tissues.

Abnormal methylation of CpG islands associated with tumor suppressor genes may also cause decreased gene expression. Increased methylation of such regions may lead to progressive reduction of normal gene expression resulting in the selection of a population of  
20 cells having a selective growth advantage (*i.e.*, a malignancy).

It is considered that altered DNA methylation patterns, particularly methylation of cytosine residues, cause genome instability and are mutagenic. This, presumably, has led to an 80% suppression of a CpG methyl acceptor site in eukaryotic organisms, which methylate their genomes. Cytosine methylation further contributes to generation of polymorphism and germ-  
25 line mutations and to transition mutations that inactivate tumor-suppressor genes (Jones, *Cancer Res.* 56:2463-2467, 1996). Methylation is also required for embryonic development of mammals (Bestor and Jaenisch, *Cell* 69:915-926, 1992). It appears that that the methylation of CpG-rich promoter regions may be blocking transcriptional activity. Therefore, there is a probability that alterations of methylation are an important epigenetic criteria and can play a  
30 role in carcinogenesis in general due to its function of regulating gene expression. Ushijima et al. (*Proc. Natl. Acad. Sci. USA* 94:2284-2289, 1997) characterized and cloned DNA fragments that show methylation changes during murine hepatocarcinogenesis. Data from a group of studies of altered methylation sites in cancer cells show that it is not simply the overall levels of DNA methylation that are altered in cancer, but changes in the distribution of methyl  
35 groups.

These studies suggest that methylation, at CpG-rich sequences known as CpG islands, provide an alternative pathway for the inactivation of tumor suppressors, despite the fact that the supporting studies have analyzed only a few restriction enzyme sites without much knowledge as to their relevance to gene control. These reports suggest that methylation of

CpG oligonucleotides in the promoters of tumor suppressor genes can lead to their inactivation. Other studies provide data that suggest that alterations in the normal methylation process are associated with genomic instability (Lengauer et al. *Proc. Natl. Acad. Sci. USA* 94:2545-2550, 1997). Such abnormal epigenetic changes may be found in many types of cancer and can, therefore, serve as potential markers for oncogenic transformation, provided that there is a reliable means for rapidly determining such epigenetic changes. The present invention was made to provide such a universal means for determining abnormal epigenetic changes and address this need in the art.

#### Methods to Determine DNA Methylation

There is a variety of genome scanning methods that have been used to identify altered methylation sites in cancer cells. For example, one method involves restriction landmark genomic scanning (Kawai et al., *Mol. Cell. Biol.* 14:7421-7427, 1994), and another example involves methylation-sensitive arbitrarily primed PCR (Gonzalzo et al., *Cancer Res.* 57:594-599, 1997). Changes in methylation patterns at specific CpG sites have been monitored by digestion of genomic DNA with methylation-sensitive restriction enzymes followed by Southern analysis of the regions of interest (digestion-Southern method). The digestion-Southern method is a straightforward method but it has inherent disadvantages in that it requires a large amount of DNA (at least or greater than 5  $\mu$ g) and has a limited scope for analysis of CpG sites (as determined by the presence of recognition sites for methylation-sensitive restriction enzymes). Another method for analyzing changes in methylation patterns involves a PCR-based process that involves digestion of genomic DNA with methylation-sensitive restriction enzymes prior to PCR amplification (Singer-Sam et al., *Nucl. Acids Res.* 18:687, 1990). However, this method has not been shown effective because of a high degree of false positive signals (methylation present) due to inefficient enzyme digestion of overamplification in a subsequent PCR reaction.

Genomic sequencing has been simplified for analysis of DNA methylation patterns and 5-methylcytosine distribution by using bisulfite treatment (Frommer et al., *Proc. Natl. Acad. Sci. USA* 89:1827-1831, 1992). Bisulfite treatment of DNA distinguishes methylated from unmethylated cytosines, but original bisulfite genomic sequencing requires large-scale sequencing of multiple plasmid clones to determine overall methylation patterns, which prevents this technique from being commercially useful for determining methylation patterns in any type of a routine diagnostic assay.

In addition, other techniques have been reported which utilize bisulfite treatment of DNA as a starting point for methylation analysis. These include methylation-specific PCR (MSP) (Herman et al. *Proc. Natl. Acad. Sci. USA* 93:9821-9826, 1992); and restriction enzyme digestion of PCR products amplified from bisulfite-converted DNA (Sadri and Hornsby, *Nucl. Acids Res.* 24:5058-5059, 1996; and Xiong and Laird, *Nucl. Acids Res.* 25:2532-2534, 1997).

PCR techniques have been developed for detection of gene mutations (Kuppuswamy et al., *Proc. Natl. Acad. Sci. USA* 88:1143-1147, 1991) and quantitation of allelic-specific

expression (Szabo and Mann, *Genes Dev.* 9:3097-3108, 1995; and Singer-Sam et al., *PCR Methods Appl.* 1:160-163, 1992). Such techniques use internal primers, which anneal to a PCR-generated template and terminate immediately 5' of the single nucleotide to be assayed. However an allelic-specific expression technique has not been tried within the context of assaying for DNA methylation patterns.

Therefore, there is a need in the art to develop improved diagnostic assays for early detection of cancer using reliable and reproducible methods for determining DNA methylation patterns that can be performed using familiar procedures suitable for widespread use. This invention was made to address the foregoing need.

### Summary of the Invention

The present invention provides a method for determining DNA methylation patterns at cytosine sites, comprising the steps of:

- (a) obtaining genomic DNA from a DNA sample to be assayed;
- (b) reacting the genomic DNA with sodium bisulfite to convert unmethylated cytosine residues to uracil residues while leaving any 5-methylcytosine residues unchanged to provide primers specific for the bisulfite-converted genomic sample for top strand or bottom strand methylation analysis;
- (c) performing a PCR amplification procedure using the top strand or bottom strand specific primers;
- (d) isolating the PCR amplification products;
- (e) performing a primer extension reaction using Ms-SNuPE primers, [<sup>32</sup>P]dNTPs and *Taq* polymerase, wherein the Ms-SNuPE primers comprise from about a 15 mer to about a 22 mer length primer that terminates immediately 5' of a single nucleotide to be assayed; and
- (f) determining the relative amount of methylation at CpG sites by measuring the incorporation of different <sup>32</sup>P-labeled dNTPs.

Preferably, the [<sup>32</sup>P]NTP for top strand analysis is [<sup>32</sup>P]dCTP or [<sup>32</sup>P]TTP. Preferably, the [<sup>32</sup>P]NTP for bottom strand analysis is [<sup>32</sup>P]dATP or [<sup>32</sup>P]dGTP. Preferably, the isolation step of the PCR products uses an electrophoresis technique. Most preferably, the electrophoresis technique uses an agarose gel. Preferably, the Ms-SNuPE primer sequence comprises a sequence of at least fifteen but no more than twenty five, bases having a sequence selected from the group consisting of GaL1 [SEQ ID NO. 1], GaL2 [SEQ ID NO. 2], GaL4 [SEQ ID NO. 3], HuN1 [SEQ ID NO. 5], HuN2 [SEQ ID NO. 6], HuN3 [SEQ ID NO. 7], HuN4 [SEQ ID NO. 8], HuN5 [SEQ ID NO. 8], HuN6 [SEQ ID NO. 9], CaS1 [SEQ ID NO. 10], CaS2 [SEQ ID NO. 11], CaS4 [SEQ ID NO. 12], and combinations thereof.

The present invention further provides a Ms-SNuPE primer sequence designed to anneal to and terminate immediately 5' of a desired cytosine codon in the CpG target site and that is located 5' upstream from a CpG island and are frequently hypermethylated in promoter regions of somatic genes in malignant tissue. Preferably, the Ms-SNuPE primer sequence

comprises a sequence of at least fifteen bases having a sequence selected from the group consisting of GaL1 [SEQ ID NO. 1], GaL2 [SEQ ID NO. 2], GaL4 [SEQ ID NO. 3], HuN1 [SEQ ID NO. 5], HuN2 [SEQ ID NO. 6], HuN3 [SEQ ID NO. 7], HuN4 [SEQ ID NO. 8], HuN5 [SEQ ID NO. 8], HuN6 [SEQ ID NO. 9], CaS1 [SEQ ID NO. 10], CaS2 [SEQ ID NO. 11], CaS4 [SEQ ID NO. 12], and combinations thereof. The present invention further provides a method for obtaining a Ms-SNuPE primer sequence, comprising finding a hypermethylated CpG island in a somatic gene from a malignant tissue or cell culture, determining the sequence located immediately 5' upstream from the hypermethylated CpG island, and isolating a 15 to 25 mer sequence 5' upstream from the hypermethylated CpG island for use as a Ms-SNuPE primer. The present invention further provides a Ms-SNuPE primer comprising a 15 to 25 mer oligonucleotide sequence obtained by the process comprising, finding a hypermethylated CpG island in a somatic gene from a malignant tissue or cell culture, determining the sequence located immediately 5' upstream from the hypermethylated CpG island, and isolating a 15 to 25 mer sequence 5' upstream from the hypermethylated CpG island for use as a Ms-SNuPE primer.

### Brief Description of the Drawings

Figure 1 shows a diagram of the inventive Ms-SNuPE assay for determination of strand-specific methylation status at cytosines. The process involves treating genomic DNA with sodium bisulfite, and generating a template by a PCR technique for a top strand methylation analysis. Alternatively a bottom strand methylation can also be assayed by designing the appropriate primers to generate a bottom strand-specific template. The process further entails amplifying the templates by a PCR technique. The PCR products are electrophoresed and isolated from agarose gels, followed by incubation with Ms-SNuPE primers, as disclosed herein wherein the Ms-SNuPE primers comprise a from about a 15 mer to about a 25 mer length primer that terminates immediately 5' of a single nucleotide to be assayed, and PCR buffer, [<sup>32</sup>P]dNTPs and *Taq* polymerase for primer extension reactions. The radiolabeled products are separated, for example, by electrophoresis on polyacrylamide gels under denaturing conditions and visualized by exposure to autoradiographic film or phosphorimage quantitation.

Figure 2 shows the results from a quantitative methylation analysis of three top strand CpG sites from a 5' CpG island of *p16*. *P16* is a known tumor suppressor gene and the particular region examined for changes in methylation is the promoter region of this gene. The top panel provides the locations of three sites analyzed (numbered 1, 2 and 3) relative to the putative transcriptional start sites (vertical arrows pointing upwards) and the exon 1α coding domain. The PCR primers used for top strand amplification of the 5' region of *p16* (which includes putative transcriptional start sites) were 5'-GTA GGT GGG GAG GAG TTT AGT T-3' [SEQ ID NO. 13] and 5'-TCT AAT AAC CAA CCA ACC CCT CC-3' [SEQ ID NO. 14]. The control sets included "M" PCR product amplified from a plasmid containing bisulfide-

specific methylated sequence; "U" PCR product amplified from a plasmid containing bisulfite-specific unmethylated sequence; and "mix" a 50:50 mixture of methylated and unmethylated PCR-amplified plasmid sequences. The DNA samples analyzed included T24 and J82 bladder cancer cell lines; wbc (white blood cell), melanoma (primary melanoma tumor tissue sample), and bladder (primary bladder tumor tissue sample). The tissue samples were micro dissected from paraffin-embedded tumor material. The grid at the bottom of the lower panel shows the ratio of methylated (C) versus unmethylated (T) bands at each site based upon phosphorimage quantitation.

Figure 3 shows a mixing experiment showing a linear response of the inventive Ms-SNuPE assay for detection of cytosine methylation. A T24 bladder cancer cell line DNA (predominantly methylated) was added in increasing amounts to a J82 bladder cancer cell line DNA (predominantly unmethylated). Figure 3 shows data from an 18 mer oligonucleotide [SEQ ID NO. 16] which was used in multiplex analysis of CpG methylation (site 2) of the *p16* 5'CpG in combination with a 15-mer and 21-mer primer [SEQ ID NOS 17 and 15, respectively] (correlation coefficient =0.99). Both the 15 mer and 21-mer produced a nearly identical linear response as the 18-mer. Figure 3 shows data from three separate experiments.

Figure 4 shows a schematic diagram that outlines a process for a high-throughput methylation analysis. The Ms-SNuPE primer extension reactions are performed and then the products are directly transferred to membranes, preferably nylon membranes. This allows for a large number of samples to be analyzed simultaneously in a high-density format. The membrane is washed and exposed to a phosphorimage cassette for quantitative methylation analysis and eliminate the need for polyacrylamide gel electrophoresis for data measurement.

Figure 5 (Panel A) shows results from quantitative analysis of DNA methylation using the Ms-SNuPE blot transfer technique of Figure 4. Levels of DNA methylation in matched normal and tumor colon specimens were analyzed in the 5' promoter region of the *p16* gene. The average methylation of 3 sites in the *p16* promoter (Figure 2) was determined by quantitating the C:T signal ration by phosphorimage analysis. Panel B shows the results of quantitating the average methylation of 3 CpG sites using standard polyacrylamide gel electrophoresis compared to dot blot transfers. The average methylation of the monitored sites in various colon specimens is plotted on the graph and shows little difference between quantitated values derived from polyacrylamide gel electrophoresis compared tom the dotblot technique. These data show the feasibility of using the Ms-SNuPE dotblot procedure for high-throughput detection and quantitation of DNA methylation changes in cancer cells.

### **Detailed Description of the Invention**

The present invention provides a method for determining DNA methylation patterns at cytosine sites, comprising the steps of:



- (a) obtaining genomic DNA from a DNA sample to be assayed, wherein sources of DNA include, for example, cell lines, blood, sputum, stool, urine, cerebrospinal fluid, paraffin-embedded tissues, histological slides and combinations thereof;
- (b) reacting the genomic DNA with sodium bisulfite to convert unmethylated  
5 cytosine residues to uracil residues while leaving any 5-methylcytosine residues unchanged to provide primers specific for the bisulfite-converted genomic sample for top strand or bottom strand methylation analysis;
- (c) performing a PCR amplification procedure using the top strand or bottom strand specific primers;
- 10 (d) isolating the PCR amplification products;
- (e) performing a primer extension reaction using Ms-SNuPE primers, [<sup>32</sup>P]dNTPs and *Taq* polymerase, wherein the Ms-SNuPE primers comprise a from about a 15 mer to about a 22 mer length primer that terminates immediately 5' of a single nucleotide to be assayed; and
- (f) determining the relative amount of allelic expression of CpG methylated sites  
15 by measuring the incorporation of different <sup>32</sup>P-labeled dNTPs.

Preferably, the [<sup>32</sup>P]NTP for top strand analysis is [<sup>32</sup>P]dCTP or [<sup>32</sup>P]TTP. Preferably, the [<sup>32</sup>P]NTP for bottom strand analysis is [<sup>32</sup>P]dATP or [<sup>32</sup>P]dGTP. Preferably, the isolation step of the PCR products uses an electrophoresis technique. Most preferably, the electrophoresis technique uses an agarose gel.

- 20 DNA is isolated by standard techniques for isolating DNA from cellular, tissue or specimen samples. Such standard methods are found in textbook references such as Fritsch and Maniatis eds., *Molecular Cloning: A Laboratory Manual*, 1989.

The bisulfite reaction is performed according to standard techniques. For example and briefly, approximately 1 microgram of genomic DNA (amount of DNA can be less when using  
25 micro-dissected DNA specimens) is denatured for 15 minutes at 45 °C with 2N NaOH followed by incubation with 0.1M hydroquinone and 3.6M sodium bisulfite (pH 5.0) at 55 °C for 12 hours (appropriate range is 4-12 hours). The DNA is then purified from the reaction mixture using standard (commercially-available) DNA miniprep columns, or other standard techniques for DNA purification are also appropriate. The purified DNA sample is  
30 resuspended in 55 microliters of water and 5 microliters of 3N NaOH is added for a desulfonation reaction, preferably performed at 40 °C for 5-10 minutes. The DNA sample is then ethanol-precipitated and washed before being resuspended in an appropriate volume of water. Bisulfite treatment of DNA distinguishes methylated from unmethylated cytosines. The present bisulfite treatment method has advantages because it is quantitative, does not use  
35 restriction enzymes, and many CpG sites can be analyzed in each primer extension reaction by using a multiplex primer strategy.

The PCR amplification step (c) can be performed by standard PCR techniques, following a manufacturer's instructions. For example, approximately 1-2 microliters of the bisulfite-treated DNA was used as a template for strand-specific PCR amplification in a region

of interest. In a PCR reaction profile for amplifying a portion of the *p16* 5' CpG island, for example, a procedure of initial denaturation of 94 °C for 3 minutes followed by a cycle of 94 °C of 30 seconds, 68 °C for 30 seconds, 72 °C for 30 seconds for a total of 30 cycles. The PCR reactions were performed in 25 microliter volumes under conditions of: ~50 ng bisulfite-converted DNA (less for micro dissected samples), 10 mM Tris-HCl (pH 8.3), 1.5 mM MgCl<sub>2</sub>, 50 mM KCl, 0.1% gelatin/ml, 100 µM of each of dNTP, 0.5 µM final concentration of each primer and 1 unit of *Taq* polymerase. There are many chromatographic techniques that can be used to isolate the PCR amplification products. In one illustrative procedure, approximately 10-25 microliters of the amplified PCR products were loaded onto 2% agarose gels and electrophoresed. The bands were visualized and isolated using standard gel purification procedures.

The primer extension reaction is conducted using standard PCR primer extension techniques but using Ms-SNuPE primers as provided herein. Approximately 10-50 nanograms of purified PCR template is used in each Ms-SNuPE reaction. A typical reaction volume is about 25 microliters and comprises PCR template (about 10-50 ng), 1X PCR buffer, 1 µM of each Ms-SNuPE primer, 1 µCi of the appropriate <sup>32</sup>P-labeled dNTP (either [<sup>32</sup>P]dCTP, [<sup>32</sup>P]TTP, [<sup>32</sup>P]dATP, [<sup>32</sup>P]dGTP or combinations thereof), and 1 unit of *Taq* polymerase. As a general rule, oligonucleotides used in the primer extension reactions were designed to have annealing temperatures within 2-3 °C of each other and did not hybridize to sequences that originally contained CpG dinucleotides. The Ms-SNuPE reactions were performed at 95 °C for 1 minute, 50 °C for 2 minutes, and 72 °C for 1 minute. A stop solution (10 microliters) was added to the mixtures to terminate the reactions. The inventive Ms-SNuPE assay utilizes internal primer(s) which anneal to a PCR-generated template and terminate immediately 5' of the single nucleotide to be assayed. A similar procedure has been used successfully for detection of gene mutations Kuppaswamy et al., *Proc. Natl. Acad. Sci. USA* 88:1143-1147, 1991) and for quantitation of allele-specific expression (Szabo and Mann, *Genes Dev.* 9:3097-3108, 1995 and Greenwood and Burke, *Genome Res.* 6:336-348, 1996).

There are several techniques that are able to determine the relative amount of methylation at each CpG site, for example, using a denaturing polyacrylamide gel to measure <sup>32</sup>P through phosphorimage analysis, or transfer of Ms-SNuPE reaction products to nylon membranes, or even using fluorescent probes instead of a <sup>32</sup>P marker. In one method for determining the relative amount of methylation at each CpG site, approximately 1-2 microliters of each Ms-SNuPE reaction product was electrophoresed onto 15% denaturing polyacrylamide gel (7M urea). The gels were transferred to filter paper and then dried. Phosphorimage analysis was performed to determine the relative amount of radiolabeled incorporation. An alternative method for determining the relative amount of methylation at individual CpG sites is by a direct transfer of the Ms-SNuPE reaction products to nylon membranes. This technique can be used to quantitate an average percent methylation of multiple CpG sites without using polyacrylamide gel electrophoresis. High-throughput methylation analysis was performed by

direct transfer of the Ms-SNuPE reactions onto nylon membranes. A total of 100 microliters or 0.4 mM NaOH, 1 mM  $\text{Na}_4\text{P}_2\text{O}_7$  was added to the completed primer extension reactions instead of adding stop solution. The mixture was directly transferred to nylon membranes using a dotblot vacuum manifold in a 96 well plate format. Each vacuum transfer well was washed a total of 4 times with 200 microliters of 2X SSC, 1 mM  $\text{Na}_4\text{P}_2\text{O}_7$ . The entire membrane was washed in 2X SSC, 1 mM  $\text{Na}_4\text{P}_2\text{O}_7$ . The radioactivity of each spot on the dried nylon membrane was quantitated by phosphorimaging analysis.

In the inventive quantitative Ms-SNuPE assay, the relative amount of allelic expression is quantitated by measuring the incorporation of different  $^{32}\text{P}$ -labeled dNTPs. Figure 1 outlines how the assay can be utilized for quantitative methylation analysis. For example, the initial treatment of genomic DNA with sodium bisulfite causes unmethylated cytosine to be converted to uracil, which is subsequently replicated as thymine during PCR. Methylcytosine is resistant to deamination and is replicated as cytosine during amplification. Quantitation of the ratio of methylated versus unmethylated cytosine (C versus T) at the original CpG sites can be determined by incubating a gel-isolated PCR product, primer(s) and *Taq* polymerase with either  $^{32}\text{P}$ -dCTP or  $^{32}\text{P}$ -TTP, followed by denaturing polyacrylamide gel electrophoresis and phosphorimage analysis. In addition, opposite strand (bottom strand) Ms-SNuPE primers are further designed which would incorporate either  $^{32}\text{P}$ -dATP or  $^{32}\text{P}$ -dGTP to assess methylation status depending on which CpG site is analyzed.

#### Ms-SNuPE Primers

The present invention further provides a Ms-SNuPE primer sequence designed to anneal to and terminate immediately 5' of a desired cytosine codon in the CpG target site and that is located 5' upstream from a CpG island and are frequently hypermethylated in promoter regions of somatic genes in malignant tissue. Preferably, the Ms-SNuPE primer sequence comprises a sequence of at least fifteen bases having a sequence selected from the group consisting of GaL1 [SEQ ID NO. 1], GaL2 [SEQ ID NO. 2], GaL4 [SEQ ID NO. 3], HuN1 [SEQ ID NO. 5], HuN2 [SEQ ID NO. 6], HuN3 [SEQ ID NO. 7], HuN4 [SEQ ID NO. 8], HuN5 [SEQ ID NO. 8], HuN6 [SEQ ID NO. 9], CaS1 [SEQ ID NO. 10], CaS2 [SEQ ID NO. 11], CaS4 [SEQ ID NO. 12], and combinations thereof. The present invention further provides a method for obtaining a Ms-SNuPE primer sequence, comprising finding a hypermethylated CpG island in a somatic gene from a malignant tissue or cell culture, determining the sequence located immediately 5' upstream from the hypermethylated CpG island, and isolating a 15 to 25 mer sequence 5' upstream from the hypermethylated CpG island for use as a Ms-SNuPE primer. The present invention further provides a Ms-SNuPE primer comprising a 15 to 25 mer oligonucleotide sequence obtained by the process comprising, (a) identifying hypermethylated CpG islands a somatic gene from a malignant tissue or cell culture source, (b) determining the sequence located immediately 5' upstream from the hypermethylated CpG island, and (c) isolating at least a 15 mer sequence 5' upstream from the hypermethylated CpG island for use

as a Ms-SNuPE primer. Preferably the Ms-SNuPE primer sequence is from about 15 to about 25 base pairs in length.

5 The ability to detect methylation changes associated with oncogenic transformation is of critical importance in understanding how DNA methylation may contribute to tumorigenesis. Regions of DNA that have tumor-specific methylation alterations can be accomplished using a variety of techniques. This will permit rapid methylation analysis of specific CpG sites using the inventive quantitative Ms-SNuPE primer process. For example, techniques such as restriction landmark genomic scanning (RLGS) (Hatada et al., *Proc. Natl.*  
10 *Acad. Sci. USA* 88:9523-9527, 1995), methylation-sensitive-representational difference analysis (MS-RDA) (Ushijima et al., *Proc. Natl. Acad. Sci. USA* 94:2284-2289, 1997) and methylation-sensitive arbitrarily primed PCR (AP-PCR) (Gonzalzo et al., *Cancer Res.* 57: 594-599, 1997) can be used for identifying and characterizing methylation differences between genomes.

15 Briefly, sequence determinations of regions of DNA that show tumor-specific methylation changes can be performed using standard techniques, such as those procedures described in textbook references such as Fritsch and Maniatis eds., *Molecular Cloning: A Laboratory Manual*, 1989. Additionally, commercially available kits or automated DNA sequencing systems can be utilized. Once specific regions of DNA have been identified by  
20 using such techniques, the Ms-SNuPE primers can be applied for rapidly screening the most important CpG sites that are involved with the specific methylation changes associated with a cancer phenotype.

### Example 1

25 This example illustrates a quantitative methylation analysis of three top strand sites in a 5' CpG island of *p16* in various DNA samples using the inventive method. The top panel provides the locations of three sites analyzed (numbered 1, 2 and 3) relative to the putative transcriptional start sites (vertical arrows pointing upwards) and the exon 1 $\alpha$  coding domain. The PCR primers used for top strand amplification of the 5' region of *p16* (which includes  
30 putative transcriptional start sites) were 5'-GTA GGT GGG GAG GAG TTT AGT T-3' [SEQ ID NO. 13] and 5'-TCT AAT AAC CAACCA ACC CCT CC-3' [SEQ ID NO. 14]. The reactions were performed in 25  $\mu$ l total volume under the conditions of 50 ng bisulfite-treated DNA, 10 mM Tris-HCl (pH 8.3), 1.5 mM MgCl<sub>2</sub>, 50 mM KCl, 0.1% gelatin/ml, 100  $\mu$ M of each dNTP, 0.5  $\mu$ M final concentration of each primer and 1U of *Taq* polymerase (Boehinger  
35 Mannheim, Indianapolis, IN). The reactions were hot-started using a 1:1 mixture of *Taq/TaqStart* antibody (Clontech, Palo Alto, CA).

An initial denaturation of 94 °C for 3 minutes was followed by 94 °C for 30 sec, 68 °C for 30 sec, 72 °C for 30 sec for a total of 35 cycles. The PCR products were separated by

electrophoresis on 2% agarose gels and the bands were isolated using a Qiaquick™ gel extraction kit (Qiagen, Santa Clarita, CA).

The Ms-SNuPE reaction was performed in a 25 ml reaction volume with 10-50 ng of PCR template incubated in a final concentration of 1 x PCR buffer, 1 μM of each Ms-SNuPE primer, 1 μCi of either [<sup>32</sup>P]dCTP or [<sup>32</sup>P]TTP and 1U of *Taq* polymerase. The primer extensions were also hot-started using a 1: mixture of *Taq/TaqStart* antibody. The primers used for the Ms-SNuPE analysis were: site 1 5'-TTT TTT TGT TTG GAA AGA TAT-3' [SEQ ID NO. 15]; site 2 5'-TTT TAG GGG TGT TAT ATT-3' [SEQ ID NO. 16]; site 3 5'-TTT GAG GGA TAG GGT-3' [SEQ ID NO. 17]. The conditions for the primer extension reactions were 95 °C for 1 minute, 50 °C for 2 minutes and 70 °C for 1 minute. A stop solution (10 μl) was added to the reaction mixtures and the samples were loaded onto 15% denaturing polyacrylamide gels (7 M urea). Radioactivity of the bands was quantitated by phosphorimaging analysis. The control sets included "M" PCR product amplified from a plasmid containing bisulfite-specific methylated sequence; "U" PCR product amplified from a plasmid containing bisulfite-specific unmethylated sequence; and "mix" a 50:50 mixture of methylated and unmethylated PCR-amplified plasmid sequences. The DNA samples analyzed included T24 and J82 bladder cancer cell lines; wbc (white blood cell), melanoma (primary melanoma tumor tissue sample), and bladder (primary bladder tumor tissue sample). The tissue samples were micro dissected from paraffin-embedded tumor material. The grid at the bottom of the lower panel shows the ratio of methylated (C) versus unmethylated (T) bands at each site based upon phosphorimage quantitation.

These data (Figure 2) show the ability of the inventive assay to detect altered patterns of methylation.

25

### Example 2

This example illustrates a mixing experiment showing a linear response of the inventive Ms-SNuPE assay for detection of cytosine methylation. A T24 bladder cancer cell line DNA (predominantly methylated) was added in increasing amounts to a J82 bladder cancer cell line DNA (predominantly unmethylated). Figure 3 shows data from an 18 mer oligonucleotide [SEQ ID NO. 16] which was used in multiplex analysis of CpG methylation (site 2) of the *p16* 5'CpG in combination with a 15-mer and 21-mer primer [SEQ ID NOS 17 and 15, respectively] (correlation coefficient =0.99). Both the 15 mer and 21-mer produced a nearly identical linear response as the 18-mer. Figure 3 shows data from three separate experiments. Differential specific activity and incorporation efficiency of each [<sup>32</sup>P]dNTP was controlled for by using a 50:50 mixture of bisulfite-specific methylated versus unmethylated PCR template for analysis.

35

### Example 3

This example provides a summary of DNA regions for which Ms-SNuPE primers can

be designed and the inventive method applied for a quantitative detection of abnormal DNA methylation in cancer cells. The sequences are listed according to name, size and frequency of hypermethylation in the corresponding cell line or primary tumor.

fragment	size (bp)	methyalted in colon cell line	hypermethylated in colon cancer	hypermethylated in bladder cancer	comments
GaL1	530	7/7 (100%)	3/7 (42%)	3/7 (42%)	GC content (0.6), observed/expected CpG (0.63)
GaL2	308	7/7 (100%)	4/5 (80%)	6/7 (85%)	GC content (0.6), observed/expected CpG (0.6)
GaL4	177	7/7 (100%)	1/2 (50%)	¾ (75%)	GC content (0.59), observed/expected CpG (0.50)
CaS1	215	4/7 (57%)	0/5 (0%)	2/7 (28%)	GC content (0.55), observed/expected CpG (0.78)
CaS2	220	4/7 (57%)	3/5 (60%)	3/7 (42%)	GC content (0.54), observed/expected CpG (0.74)
CaS4	196	6/7 (85%)	0/5 (0%)	1/7 (14%)	GC content (0.64), observed/expected CpG (0.84)
HuN1	148	7/7 (100%)	3/5 (60%)	3/7 (42%)	GC content (0.54), observed/expected CpG (0.99)
HuN2	384	7/7 (100%)	4/5 (80%)	2/7 (28%)	GC content (0.6), observed/expected CpG (0.62)
HuN3	178	6/7 (85%)	4/5 (80%)	3/7 (42%)	GC content (0.53), observed/expected CpG (0.97)
HuN4	359	7/7 (100%)	3/5 (60%)	4/7 (57%)	GC content (0.51), observed/expected CpG (0.47)

HuN5	251	7/7 (100%)	2/5 (40%)	5/7 (71%)	GC content (0.63), observed/expected CpG (0.77)
HuN6	145	6/7 (85%)	3/4 (75%)	1/2 (50%)	GC content (0.55), observed/expected CpG (0.47)

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- 5 (i) APPLICANTS: Mark L. Gonzalgo and Peter A. Jones
- (ii) TITLE OF INVENTION: A CANCER DIAGNOSTIC METHOD BASED  
UPON DNA METHYLATION DIFFERENCES
- 10 (iii) NUMBER OF SEQUENCES: 17
- (iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: DAVIS WRIGHT TREMAINE
- 15 (B) STREET: 2600 Century Square, 1501 Fourth Avenue
- (C) CITY: Seattle
- (D) STATE: Washington
- (E) COUNTRY: U.S.A.
- (F) ZIP: 98101
- 20 (v) COMPUTER READABLE FORM:
- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: PC compatible
- (C) OPERATING SYSTEM: Windows95
- 25 (D) SOFTWARE: Word
- (vi) CURRENT APPLICATION DATA:
- (A) APPLICATION NUMBER: to be assigned
- (B) FILING DATE:
- 30 (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
- (A) NAME: Oster, Jeffrey B.
- (B) REGISTRATION NUMBER: 32,585
- 35 (C) REFERENCE/DOCKET NUMBER: 47675-2
- (ix) TELECOMMUNICATION INFORMATION:
- (A) TELEPHONE: 206 628 7711
- (B) TELEFAX: 206 628 7699
- 40 (2) INFORMATION FOR SEQ ID NO:1:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 530
- (B) TYPE: nucleic acid
- 45 (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: GaL1
- 50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

1 CCCGCGACCT AAGCCAGCGA CTTACCACGT TAGTCAGCTA AGAAGTGGCA 50  
51 GAGCTGGGAT TCGAACCTAT AAAGAACTCT GAAGCCTGGG TATTTTACCA 100  
55 101 TGACACTTTA CATAATGCGC CACGGGGTAG TCGGAGGGGG AGGTCCATCT 150  
151 CCCTTTCCCT TGCTGTCCAT CTCCACAGAA AAGAAGCAAG TGGAGGACAG 200  
201 GAGCCAGAAA GTCATCTGGC CGCGGATCAT TCCGGAGTGA CCCCCGCCGC 250  
251 CACCACTCGC ATAGTCCGCT TATGGCGGGA GGGCACCTCA GAGATTCTCA 300



301 CAGGGGCTGT GCGGCCAGAA CCAGAAGTGC AAAGCACCGT TAGCGACTCT 350  
351 ATCGCCCCCT GCCGCCTGTG GCGCCCAGTC CGAAGCTGCT GTTTTCAGGA 400  
401 GGGCTAGTGG GCTAAGAAAA GAGCTCACCG ACTGACTGCC CAACAGCTGT 450  
451 TGCGAGCCAG TGCTAGGCTG CAGACAGCCT TGCCAAATGT GGTGACATAA 500  
5 501 GCGGGAGGGG GGAACATTTA GAGAGCCCTA 530

(2) INFORMATION FOR SEQ ID NO:2:

10 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 308  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: unknown

15 (ii) MOLECULE TYPE: GaL2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

20 1 CTAGGGTAGG CTGGTCTGTG CTGGATACGC GTGTTCTTCT GCGGAGTTAA 50  
51 AGGGTCGGGG ACGGGGGTTC TGGACTTACC AGAGCAATTC CAGCCGGTGG 100  
101 GCGTTTGACA GCCACTTAAG GAGGTAGGGA AAGCGAGCTT CACCGGGCGG 150  
151 GCTACGATGA GTAGCATGAC GGGCAGCAGC AGCAGCAGCC AGCAAAAGCC 200  
201 TAGCAAAGTG TCCAGCTGCT GCACTGCCGC GGGGACTCCC ACATCACCAT 250  
25 251 GACTAGTTGT GCAACTCTGC AGCAGAAACG GCTTCCGAGG AACACAGGAT 300  
301 CGCGGGGG 308

(2) INFORMATION FOR SEQ ID NO:3:

30 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 177  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
35 (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: GaL4

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

40 1 GCTTCCTTTT TCTCGGCTTT CCTCACTATC CTCTCCCTGT TCGAGAGTAT 50  
51 CTCCACCAGC ACCGAGCCTC ACACGGGCTG TGCCTCCATC TTTGGAATGC 100  
101 CTACCCTTCT TTCTTGCGAA GCCCCTCCCA GGGCCAGCCC TTGTGCACCG 150  
45 151 GCTCAAGGGG ACTGCTCTCC TGCCTCG 177

(2) INFORMATION FOR SEQ ID NO:4:

50 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 148  
(B) TYPE: oligonucleotide  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: unknown

55 (ii) MOLECULE TYPE: HuN1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

1 TTGCGCCGAT CGTCAAGAAC CTCTCATCCC TGGCAGCAGC AAAGCCAATA 50  
51 TATTTCCATT TCTTATTTCA GTTTGCCACC AAAACAAAGC TGCGCGCGGC 100  
101 TGAGGGCAGG AAGGCGCTGA GACCGACCGA GAAGAAGGGA CGTCCCGG 148

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 384  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: HuN2 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

1 CAGGCCCCGCC GAGACTCCAC TCCAAC TACC AGGAAATTTT CCGTGGAGCT 50  
51 TCAATTCCTG GGACCCTCCT ACTGCGGGGA GAGTGGTTTC CCTGCCCCAC 100  
101 ACCATGCCCT AGGCCCCGAGT CTGCGGCTCT TGGGGGATCT CTCCGAGCTC 150  
151 CGACACCGTG TTCGGACCGG GTGCGCCCTG CCGCTGGGGC TCAAGCCTGC 200  
201 AGGCGTGAGA ACCGGGGGAC TCTCTATGGC ACCAAGAGCT TCACCGTGAG 250  
251 CGTAGGCAGA AGCTTCGCTT TGATCCTAGG GCTTACAAAG TCCTCCTTTG 300  
301 GCTGCCCATG ATGGTAAAG GGCAGTTGCT CACAAAGCGC GAGTGTGTGT 350  
351 GCCAGACAGT GTAAATGAGT GTTGGGACCG GCGT 384

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 178  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: HuN3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

1 GGGTCCGTTT GTGAATGCAT GAGCAGGGTG TGAGCGCCAG GGGGTTACAC 50  
51 TTCTCACGGG TTAAAACCCA GACAACTTCA CGAGGGAACC ACGTGCCATT 100  
101 TTAACAGCGT ACGGTCGGGA TCGTGGGACG TCATTAAACG GAGTGGGTTG 150  
151 AGTATGTGAC TCTGTCACCC ATTTTCTG

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 359  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: HuN4 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

1 CCCCGCGGGG CAGAATCCAA GTGAGTCAGA CACATTGCTC CCTCCCTGCT 50  
51 GCTGCCAGTC CATCTCTTTG CCAACAAACC TGCTTAAAT GCCAAAGCTG 100  
101 GTCCAAAGTT TCAGGAAAC AACTTCCGCC AGAGGGCACG TAGAGGGCAC 150

151 AGATGCTATA GATGCTTCTC TGACAAACAC TCCTGACCCC CTTGACAGAT 200  
 201 TGGAAAATAC ATGGTTCAGA AAGGGTGAGA GATTTCAACT TGAGAAGTGA 250  
 251 AACTAGGAAA AGATGGAAGG TGTCCGGATT TCTAGCTCAA GTCCACACAC 300  
 301 TGCTTCTGCT GCGGTGACTA AATCGTGGCT GTGTTCTCAT CACCTGCCTC 350  
 5 351 GCGGCGCGC 359

(2) INFORMATION FOR SEQ ID NO:8:

10 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 251
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown

15

(ii) MOLECULE TYPE: HuN5 primer

1 GCGGGGCCTG GGCACCGCGG AGGGGGGGCT TTTCTGCGCC CGGCGAAGCG 50  
 51 TGGAACTTGC GCCCTGAGGC AGCGCGGCGA GACCAGTCCA GAGACCGGGG 100  
 20 101 CGAGCCTCCT CAGGATTCTT CGCCCCAGTG CAGATGCTGT GAGCTTAGAC 150  
 151 GAGGACAGGG CATGGCACTC GGCTTGCCCC GTAGTGGACG GTGTTTTTGC 200  
 201 AGTCATGAAC CCAAACGCCG CAAACCTTGA CCGTTTCCCC ACCCGTGTTG T  
 251

25 (2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 145
- (B) TYPE: nucleic acid
- 30 (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: HuN6 primer

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

1 TGAGAGCAGC ATCCTCCCCT GCGTGTGGTT CTCTAACTTA CCTCCTGTAT 50  
 51 GGGGTCTGCG GACCCAGCAC ACCTCCCGGG CCCCCAAAAA ATTCCAGCTC 100  
 101 AAGAGCCCTA AAAATCCTTA CCCTGNNAAG GTTTGAGCTT CTCCC 145

40

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 215
- 45 (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: CaS1 primer

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

1 ACGCCGGCCA CAGTTCTTCA GTGAAACGCT TCACTCTCTG GTCATAGAGG 50  
 51 TAGGAAACTA TAGCTGTCCC AACTAAATGT CAGGACGAAT TAGCCAGCT 100  
 55 101 GGTCACGCTC ACAGTCACCG CCTCCACCAG ACTGAGCGAC CCTCCCAACG 150  
 151 GGGTTTGCCG TGTTGGGAGG ACAGCGGAGT TTCGTTGCTG TGTC AATTTG 200  
 201 TG TAGACGCG GCTGC 215

## (2) INFORMATION FOR SEQ ID NO:11:

## (i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 220 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: CaS2 primer

10  
1 CTGCTCTCTT CTCTTCTTTT CCCCTTTCCT CTCCTCTCCC TTTCCTCAGG 50  
51 TCACAGCGGA GTGAATCAGC TCGTGGTGT CTTTGTCAAC GGGCGGCCAC 100  
101 TGCCGGACTC CACCCGGCAG AAGATTGTAG AGCTAGCTCA CAGCGGGGCC 150  
151 CGGCCGTGCG ACATTTCCTG AATTCTGCAG GTGATCCTCC CGGCGCCGCC 200  
15 201 CCACTCGCCG CCCCCGCGGC 220

## (2) INFORMATION FOR SEQ ID NO:12:

## (i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 196  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: CaS4 primer

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

30 1 GGGCGGCACG GAGGGAGTCA GGAGTGAGCC CGAAGATGGA GAGAAGTCGA 50  
51 TTCGCCCAGA GAACGCAAGA CGGTGGATCA GAGATGAGTC CCAGGAACCT 100  
101 CAGAGAGCGA GGCTGACAGG CCCGGGGAGA GGACCGGGCA GGGACAAACC 150  
151 AGCGGACAGA GCAGAGCGCG AAATGGTTGA GACCGGGAAG CGACCT 196

## (2) INFORMATION FOR SEQ ID NO:13:

## (i) SEQUENCE CHARACTERISTICS:

- 40 (A) LENGTH: 22  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: Ms-SNuPE primer from p16 promoter region

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

1 5'-GTA GGT GGG GAG GAG TTT AGT T-3' 22

## (2) INFORMATION FOR SEQ ID NO:14:

## (i) SEQUENCE CHARACTERISTICS:

- 55 (A) LENGTH: 23  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: Ms-SNuPE primer from *p16* promoter region

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

5

1 5'-TCT AAT AAC CAA CCA ACC CCT CC-3

(2) INFORMATION FOR SEQ ID NO:15:

10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

15

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: Ms-SNuPE primer from *p16* promoter region

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

1 5'-TTT TTT TGT TTG GAA AGA TAT-3' 21

25 (2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18

(B) TYPE: nucleic acid

30

(C) STRANDEDNESS: single

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: Ms-SNuPE primer from *p16* promoter region

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

1 5'-TTT TAG GGG TGT TAT ATT-3' 18

40 (2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 15

(B) TYPE: nucleic acid

45

(C) STRANDEDNESS: single

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: Ms-SNuPE primer from *p16* promoter region

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

1 5'-TTT GAG GGA TAG GGT-3' 15

55

We claim:

1. A method for determining DNA methylation patterns at cytosine sites, comprising the steps of:
  - (a) obtaining genomic DNA from a DNA sample to be assayed;
  - 5 (b) reacting the genomic DNA with sodium bisulfite to convert unmethylated cytosine residues to uracil residues while leaving any 5-methylcytosine residues unchanged to provide primers specific for the bisulfite-converted genomic sample for top strand or bottom strand methylation analysis;
  - (c) performing a PCR amplification procedure using the top strand or bottom strand  
10 specific primers;
  - (d) isolating the PCR amplification products;
  - (e) performing a primer extension reaction using Ms-SNuPE primers, [<sup>32</sup>P]dNTPs and *Taq* polymerase, wherein the Ms-SNuPE primers comprise a from about a 15 mer to about a 22 mer length primer that terminates immediately 5' of a single nucleotide to be assayed; and  
15 (f) determining the relative amount of allelic expression of CpG methylated sites by measuring the incorporation of different <sup>32</sup>P-labeled dNTPs.
2. The method of claim 1 wherein the [<sup>32</sup>P]dNTP for top strand analysis is [<sup>32</sup>P]dCTP or [<sup>32</sup>P]TTP.
3. The method of claim 1 wherein the [<sup>32</sup>P]dNTP for bottom strand analysis is  
20 [<sup>32</sup>P]dATP or [<sup>32</sup>P]dGTP.
4. The method of claim 1 wherein the isolation step of the PCR products uses an electrophoresis technique.
5. The method of claim 4 wherein the electrophoresis technique uses an agarose gel.
- 25 6. The method of claim 1 wherein the Ms-SNuPE primer sequence comprises a sequence of at least fifteen but no more than twenty five bases having a sequence selected from the group consisting of GaL1 [SEQ ID NO. 1], GaL2 [SEQ ID NO. 2], GaL4 [SEQ ID NO. 3], HuN1 [SEQ ID NO. 5], HuN2 [SEQ ID NO. 6], HuN3 [SEQ ID NO. 7], HuN4 [SEQ ID NO. 8], HuN5 [SEQ ID NO. 8], HuN6 [SEQ ID NO. 9], CaS1 [SEQ ID NO. 10], CaS2 [SEQ ID  
30 NO. 11], CaS4 [SEQ ID NO. 12], and combinations thereof.
7. A Ms-SNuPE primer sequence designed to anneal to and terminate immediately 5' of a desired cytosine codon in a CpG target site, comprising an oligonucleotide sequence of at least 15 base pairs and corresponding to a gene sequence located immediately 5' upstream from the CpG island that is frequently hypermethylated in promoter regions of somatic genes  
35 in malignant tissue.
8. The Ms-SNuPE primer sequence wherein the primer sequence is from about 15 to about 25 base pairs in length and selected from the group consisting of GaL1 [SEQ ID NO. 1], GaL2 [SEQ ID NO. 2], GaL4 [SEQ ID NO. 3], HuN1 [SEQ ID NO. 5], HuN2 [SEQ ID NO. 6], HuN3 [SEQ ID NO. 7], HuN4 [SEQ ID NO. 8], HuN5 [SEQ ID NO. 8], HuN6 [SEQ

ID NO. 9], CaS1 [SEQ ID NO. 10], CaS2 [SEQ ID NO. 11], CaS4 [SEQ ID NO. 12], and combinations thereof.

- 5           9.       A method for obtaining a Ms-SNuPE primer sequence designed to anneal to and terminate immediately 5' of a desired cytosine codon in the CpG target site, comprising finding a hypermethylated CpG island in a somatic gene from a malignant tissue or cell culture, determining the sequence located immediately 5' upstream from the hypermethylated CpG island, and isolating a 15 to 25 mer sequence 5' upstream from the hypermethylated CpG island for use as a Ms-SNuPE primer.

- 10          10.       A Ms-SNuPE primer comprising a 15 to 25 mer oligonucleotide sequence obtained by the process comprising, finding a hypermethylated CpG island in a somatic gene from a malignant tissue or cell culture, determining the sequence located immediately 5' upstream from the hypermethylated CpG island, and isolating a 15 to 25 mer sequence 5' upstream from the hypermethylated CpG island for use as a Ms-SNuPE primer.

1/5

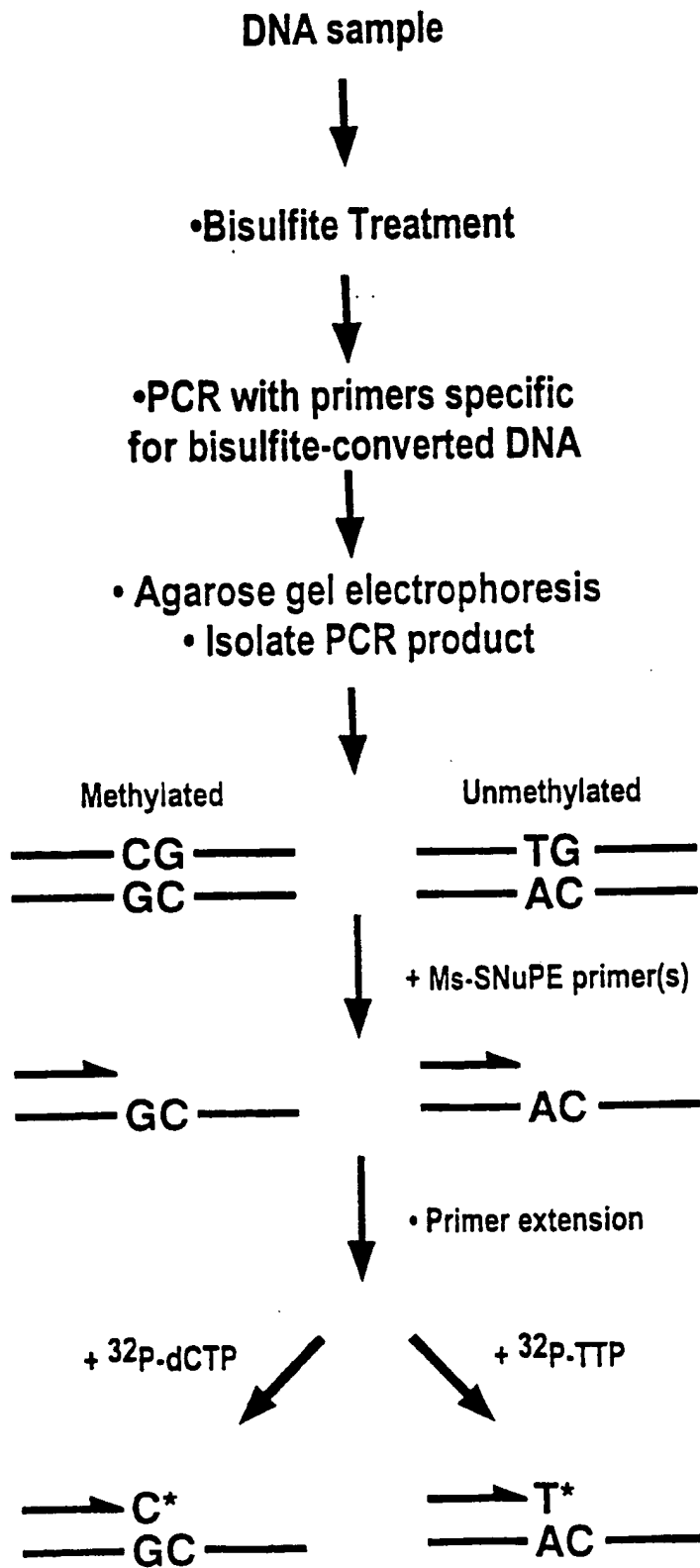


Figure 1



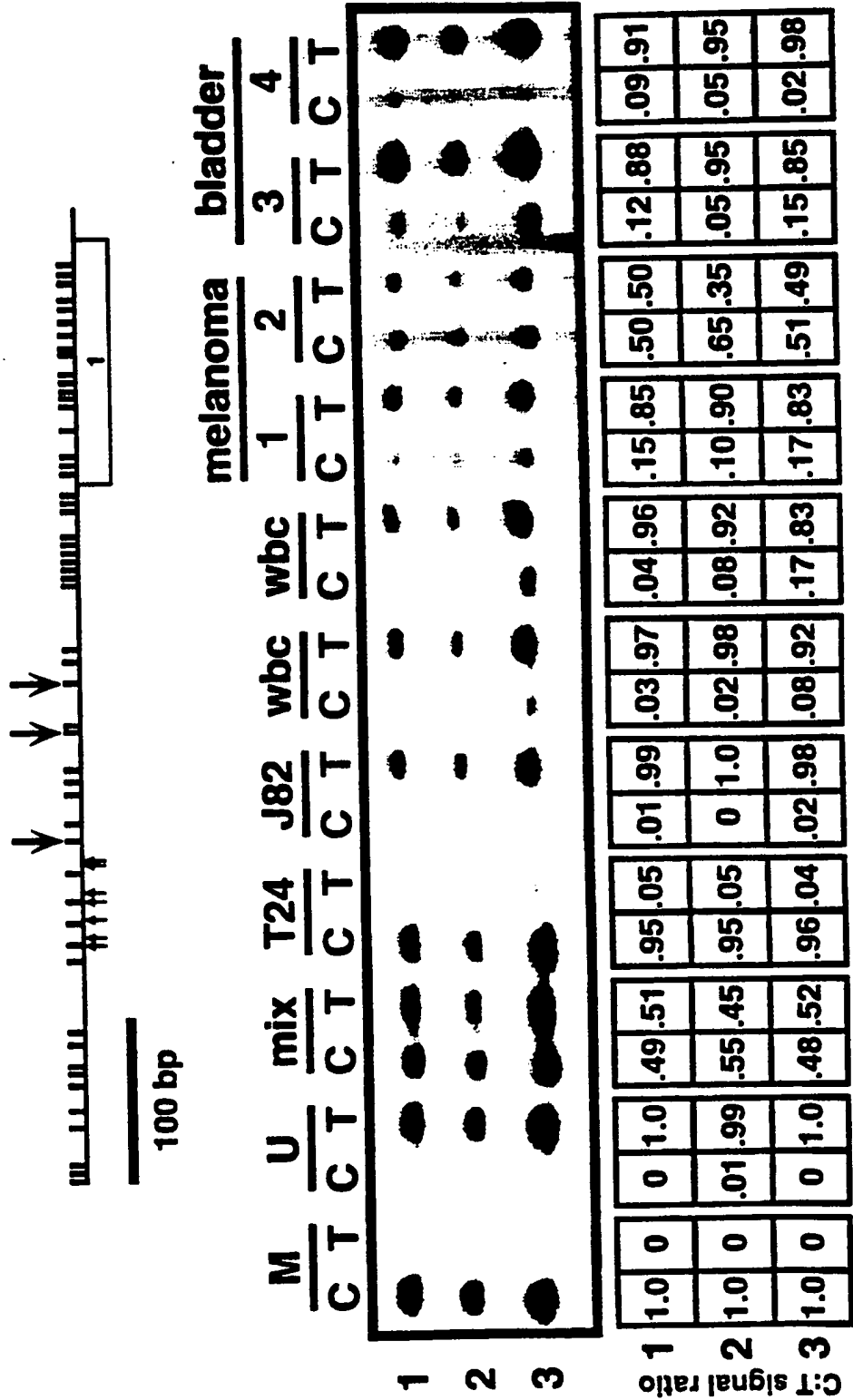


Figure 2

3/5

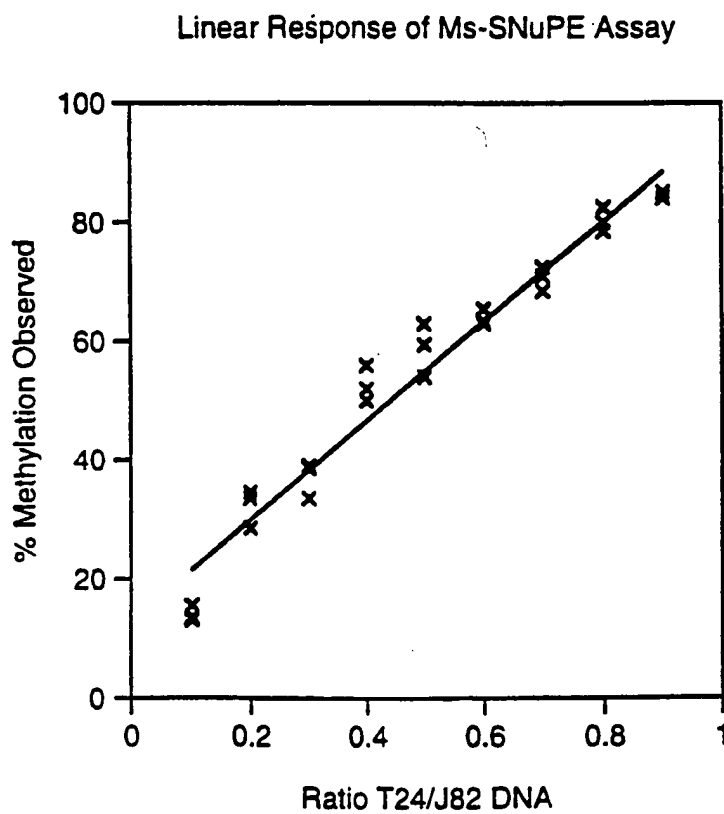


Figure 3

## High-throughput Quantitation of DNA Methylation Using Ms-SNuPE Blot Transfer Protocol

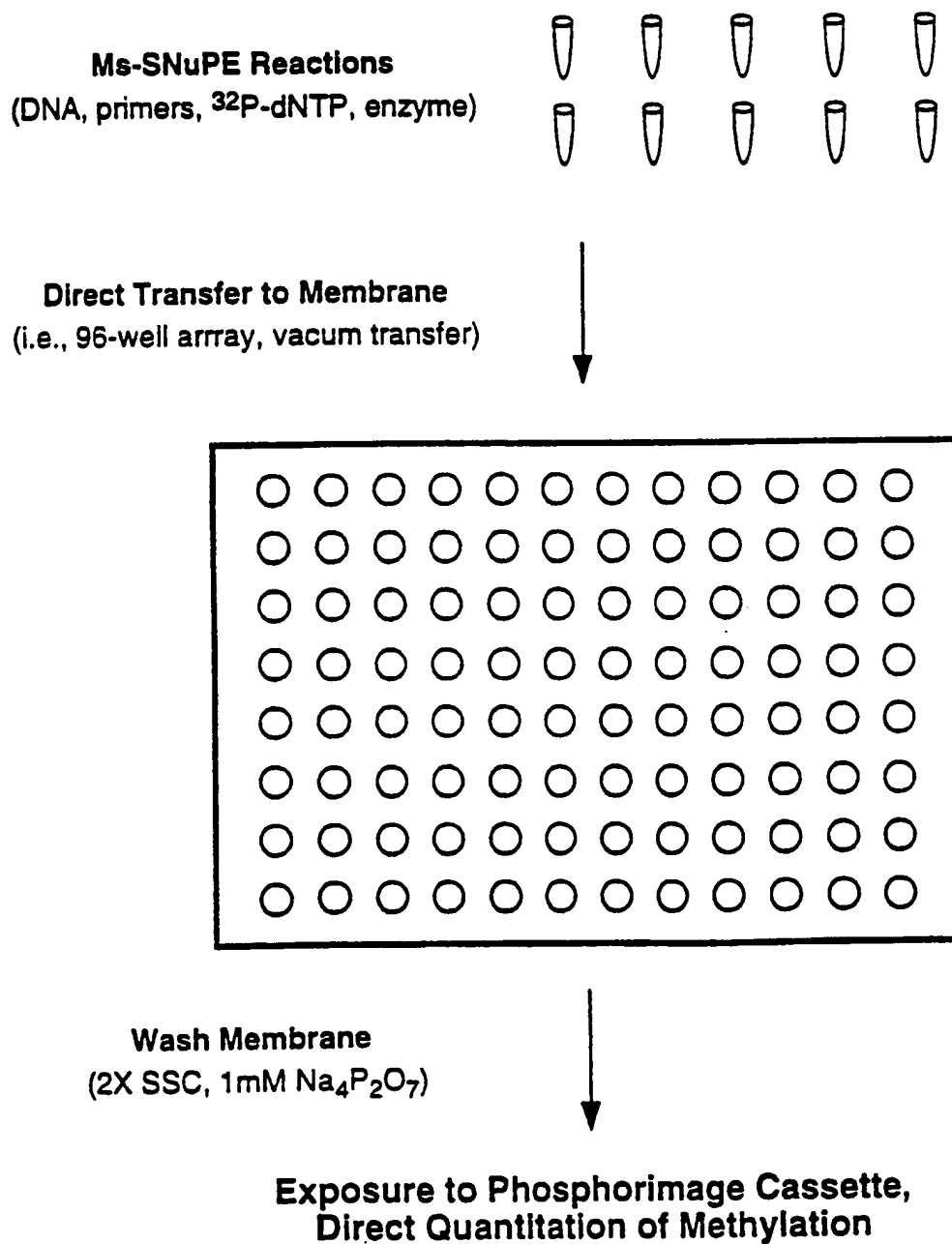
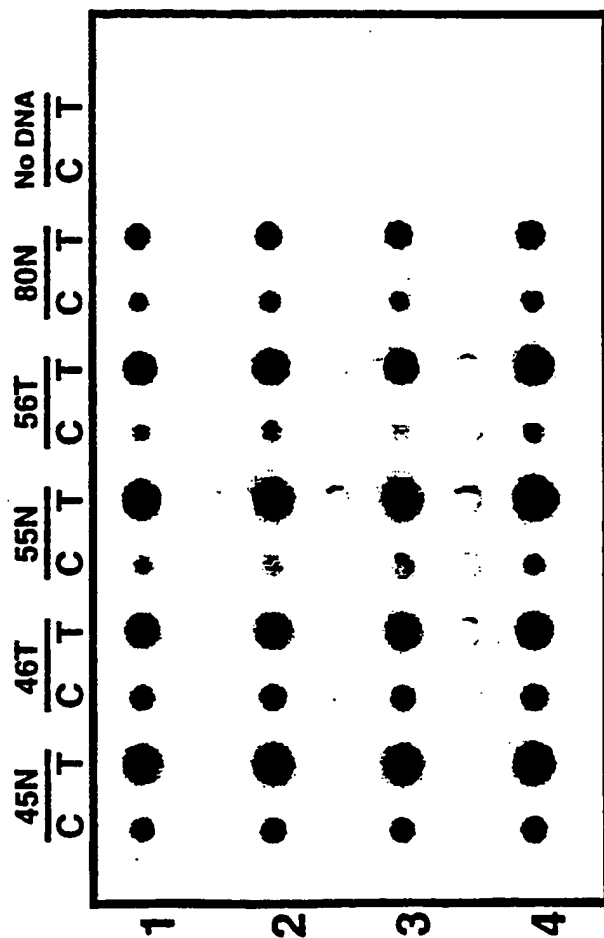


Figure 4

**A**



**B**

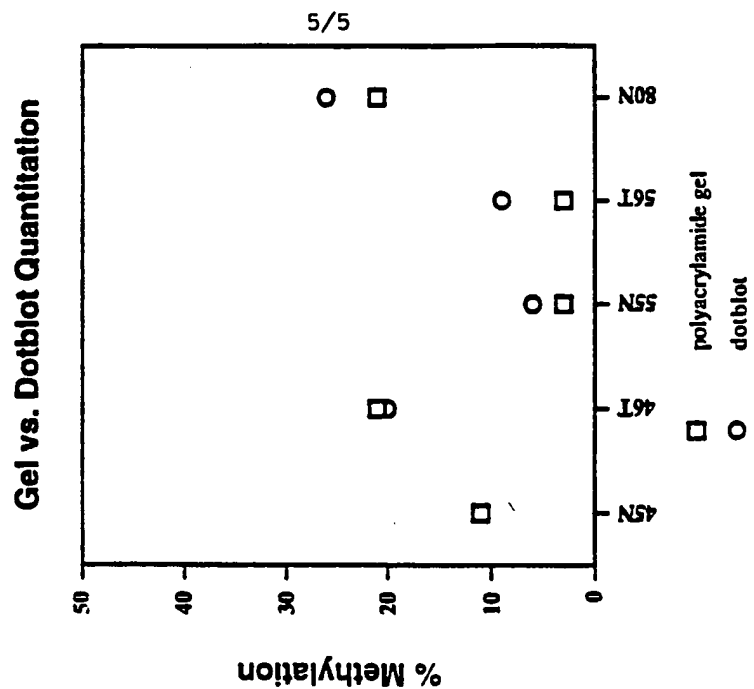


Figure 5

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US98/11896

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C12Q 1/68; C07H 21/04

US CL :435/6, 91.2; 536/24.3, 24.33

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/6, 91.2; 536/24.3, 24.33

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS; DIALOG: Medline, CA, Derwent Patents, Biosis, Embase, Scisearch  
search terms: methylation, CpG, sodium bisulfite, PCR, primer

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y, P	GONZALGO et al. Rapid quantitation of methylation differences at specific sites using methylation-sensitive single nucleotide primer extension (Ms-SNuPE). Nucleic Acids Research. 15 June 1997, Vol. 25, No. 12, pages 2529-2531, see especially Figures 1 and 2.	1-5, 7, 9, 10
A	ZHOU et al. Use of a single sequencing termination reaction to distinguish between cytosine and 5-methylcytosine in bisulfite-modified DNA. BioTechniques. May 1997, Vol. 22, No. 5, pages 850-853, especially pages 851-852.	1-5, 7, 9, 10



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A* document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*B* earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*G* document member of the same patent family
*O* document referring to an oral disclosure, use, exhibition or other means	
*P* document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

05 AUGUST 1998

Date of mailing of the international search report

21 SEP 1998

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks

Authorized officer

D. K. F. / on

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US98/11896

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2. ☒ Claims Nos.: 6 and 8  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:  
  
The claims are unsearchable to the extent that they reference specific nucleotide sequences from the Sequence Listing. Because the computer readable copy of the Sequence Listing contains errors, no meaningful search of the sequences can be carried out by this Authority.
  
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
  
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
  
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐

The additional search fees were accompanied by the applicant's protest.

☐

No protest accompanied the payment of additional search fees.

**INTERNATIONAL SEARCH REPORT**

International application No.  
PCT/US98/11896

**C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	HERMAN et al. Methylation-specific PCR: A novel PCR assay for methylation status of CpG islands. Proceedings of the National Academy of Sciences. September 1996, Vol. 93, pages 9821-9826, especially page 9822.	1-5, 7, 9, 10